

Atmos. Meas. Tech. Discuss., referee comment RC3
<https://doi.org/10.5194/amt-2021-383-RC3>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on amt-2021-383

Anonymous Referee #3

Referee comment on "Testing the efficacy of atmospheric boundary layer height detection algorithms using uncrewed aircraft system data from MOSAiC" by Gina Jozef et al., Atmos. Meas. Tech. Discuss., <https://doi.org/10.5194/amt-2021-383-RC3>, 2022

Review of « Testing the efficacy of atmospheric boundary layer height detection algorithms using uncrewed aircraft system data from MOSAiC » by Jozef et al.

This study compares different methods to determine the ABL height from uncrewed aircraft system and radio-sounding from an icebreaker in the central Arctic Ocean during the MOSAiC expedition. "Subjective" visual detection methods are compared to four "objective" methods for both the UAS and RS profiles. The difficulties of ABLH estimation largely depends on the peculiarities of ABLH in Arctic, which are shallow ABLH and a high percentage of NBL cases (and a low percentage of CBL cases).

Major comments:

- The methodology sections (§2.3-2-5) is very long and contains many redundancies. I encourage the authors to structure the paper with greater caution. For example,
 - 2.1 describes the DH2 UAS data. § L152-162 relates however to the ABL identification methods.
 - the four objective methods are cited four times (L23, L89, L156 and L300) before to be described under §2.4.
 - Part of the subjective method refers to some objective method, so that I would first describe the objective methods and then refer to similarity of the subjective method. For example the bulk Richardson method is to some extent describes at several places
 - The description of the ABL structure and the corresponding profiles -with Stull as reference (30 mentions)- is dispatched in the whole paper. A clear description of the ABL structure in the Arctic ocean would benefit the comprehension of the reader

- The term “subjective” and “objective” methods seem to be inappropriate. The “subjective method” relates to “manually” or “visual inspection” performed by a person. The objective method relates to automatic algorithms performed by computers. Moreover, the subjective method contains to some extent elements of the objective method (e.g. the use of the Rib profile, but with a threshold described as 0.25, CBLH is the parcel method, the use of RH gradient,...).
- The subjective method is considered as the best ABLH estimate. It is however also prone to error and several experiences have shown a clear ABLH uncertainty if several persons were estimating the ABLH. Were all the ABLH estimated by a unique person? A comparison between several subjective estimation could be performed? Second the criteria/profiles mostly used for the subjective SBL and NBL heights estimation are not mentioned. For example, the best correlation between the subjective method and the bulk Richardson number method found (Figs 9-10) is perhaps due to large weight of the Rib profiles in the subjective method. The subjective method is then a mixing of several criteria, a mixing that could also be applied to the objective methods. Since this exercise can be quite difficult, why not comparing visual and automatic detection with the same criteria/profiles (e.g. only with the potential temperature or the Rib profiles?)
- Similarly, the subjective method uses RH profiles, why not applying an objective gradient detection to the RH profiles?
- This comparison of the objective methods between themselves would also be of high interest and is really lacking.
- The stability regime could also be further used to explain the differences between the applied ABLH detection. For example, the symbols of Fig. 9 (apart from the blue and red colors for UAS and RS) could be differentiated as a function of the stability regime, allowing to identify potential systematic causes for the observed differences. Some similar uses could also be thought for Fig. 10.
- To avoid confusion, I would use the term “inversion layer” instead of only “inversion” to describe the atmospheric layer with temperature or humidity increase with altitude.

Minor comments:

- L63-64 and Table 1: I think that Table 1 mixes measured profiles, computed variables and, to some extent, detection methods. E.g. the virtual potential temperature is a variable that is used alone in the Parcel Method (PM) but it is also the main component of the bulk Richardson number (Rib) method. In that sense, I think that the name of the method and the used profile data are much more useful than the “variables” given in Table 1. The potential temperature is also used for the PM, Rib and gradient method. The difference between “component-wise wind speed perturbations” and “wind shear” are not directly understandable. Moreover, this list of “variables previously used to identify ABLH” is not complete (e.g. all the methods based on aerosol concentration are not mentioned).
- L69-71: rephrase, grammatical problem.
- L 74: “to either decrease or increase more above the ABL”: does this correspond to an increase/decrease of the gradient of the humidity?
- L86-95: (see also main comments) I wonder if the use of subjective/objective methods is the right one or if it corresponds to an opposition between “manual” (=performed by a person) and automatic/operational detection method. I suppose that a person

estimate the ABL heights “visually identified through combined evaluation of θ_v , humidity (both relative humidity (RH) and mixing ratio), and Rib profiles” via the same criteria as the objective methods.

- L154-159: This information has already be given.
- L164-167: Before to compare ABL height from radiosondes and UAS measurements, a comparison of the measured T, wind and humidity profiles should be performed. However, this is perhaps already included in other joined papers on the MOSAiC expedition.
- L 191: “in either direction” means in increasing/decreasing direction? I think it should be rephrased.
- L 193: does this number also depends on the uncertainties/noise of the measurements?
- L197: “between a SBL, NBL or CBL” in arctic (I suppose).
- L194-203: this could be more efficiently explained and, consequently, shortened.
- L215-216: what is meant by “ the θ_v inversion is at its strongest”? does it mean that the positive θ_v gradient is at its strongest ?
- L236: delta represent the difference between the elevation z (at which Rib is computed) and the ground level.
- L247: I don’t understand what you mean by “over 30 m bins”?
- L251-252: this is the description of the parcel method and has nothing subjective. The simultaneous increase of Rib is obvious since the bulk Richardson method with a threshold=0 corresponds to the parcel method.
- L252 “this will not be the first altitude at which the virtual potential temperature increases with altitude”: unnecessary, this is obvious.
- L 334: why is Liu and Liang’s threshold is inappropriate for the current case? Due to arctic condition? How did you identify that it is inappropriate?
- L446-447: I do not see the use of this sentence
- L467-469 should be rephrased.
- 8: It’s very nice to see the differences in ABL heights as a function of the time between launches. Other criteria such as the stability classes could also be used. Fig. 9 f) with ABLH from RS versus from DH2 also brings a nice overview of the comparison and should be discussed with Fig. 8. By the way, isn’t Figure 8 already a result so that it should appears under §3?
- L549-551: is it due to the fact that RS cannot be used below 30 m? Or due to the interpolation if the RS go through icebreaker’s plume?
- L565-570: The main point with the comparison with Liu-Liang method is 1) (as described in the manuscript) this method works well for ~40% (within 20%) of the cases and 2) in about 40% of the cases, Liu-Liang has more than 100% difference with the subjective method.
- Liu-Liang method results in the largest differences with the subjective method: is it due to a bad classification of the stability leading to a false applied detection method? Which are the stability cases with the greatest differences?